NUCLEAR MEDICINE

Can pretreatment ^{18}F-FDG PET tumor texture features predict the outcomes of osteosarcoma treated by neoadjuvant chemotherapy?

Hongjun Song[1] · Yining Jiao[2] · Weijun Wei[1,3,4] · Xuhua Ren[2,5,6] · Chentian Shen[1] · Zhongling Qiu[1] · Qingcheng Yang[7] · Qian Wang[2] · Quan-Yong Luo[1]

Abstract

Purpose To investigate whether tumor texture features derived from pretreatment with ^{18}F-fluorodeoxyglucose positron emission tomography (FDG PET) can predict histological response or event-free survival (EFS) in patients with localized osteosarcoma of the extremities treated by neoadjuvant chemotherapy (NAC).

Methods We retrospectively reviewed 35 patients with American Joint Committee on Cancer stage II extremity osteosarcoma treated with NAC and surgery. Primary tumor traditional parameters and texture features were measured for all ^{18}F-FDG PET images prior to treatment. After surgery, histological responses to NAC were evaluated on the postsurgical specimens. A receiver operating characteristic curve (ROC) was constructed to evaluate the optimal predictive performance among the various indices. EFS was calculated using the Kaplan-Meier method and prognostic significance was assessed by Cox proportional hazards analysis.

Results Pathologic examination revealed 16 (45.71%) good responders and 19 (54.29%) poor responders. Although both the texture features (least axis, dependence nonuniformity, run length nonuniformity, and size zone nonuniformity) and metabolic tumor volume (MTV) can predict tumor response of osteosarcoma to NAC, the traditional indicator MTV has the best performance according to ROC curve analysis (area under the curve = 0.918, $p < 0.0001$). In multivariate analysis, MTV ($p < 0.0001$), histological response ($p = 0.0003$), and texture feature of coarseness$_{NGTDM}$ (neighboring gray tone difference matrix) ($p = 0.005$) were independently associated with EFS.

Conclusions Intratumoral heterogeneity of baseline ^{18}F-FDG uptake measured by PET texture analysis can predict tumor response and EFS of patients with extremity osteosarcoma treated by NAC, but the conventional parameter MTV provides better predictive power and is a strong independent prognostic factor.

Key Points

• The baseline ^{18}F-FDG PET tumor texture features can predict tumor NAC response for patients with osteosarcoma.
• Coarseness$_{NGTDM}$ is a new and independent prognostic factor for osteosarcoma.
• MTV provides the best predictive power and is a strong independent prognostic factor for patients with osteosarcoma.

✉ Quan-Yong Luo
lqyn@sh163.net

Qian Wang
wang.qian@sjtu.edu.cn

1 Department of Nuclear Medicine, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, 600 Yishan Road, Shanghai 200233, People's Republic of China

2 Institute for Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, People's Republic of China

3 Department of Radiology, University of Wisconsin, Madison, WI 53705, USA

4 Department of Medical Physics, University of Wisconsin, Madison, WI 53705, USA

5 Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

6 BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

7 Department of Orthopedics, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, People's Republic of China

Abbreviations

AJCC	American Joint Committee on Cancer
AUC	Receiver operating characteristic (ROC) curves
DNU	Dependence nonuniformity
DNUN	Dependence nonuniformity normalized
DV	Dependence variance
EFS	Event-free survival
FDG	^{18}F-fluorodeoxyglucose
GLCM	Gray level co-occurrence matrix
GLDM	Gray level dependence matrix
GLNU	Gray level nonuniformity
GLRLM	Gray level run length matrix
GLSZM	Gray level size zone matrix
GLV	Gray level variance
HGLE	High gray level emphasis
IV	Inverse variance
JE	Joint entropy
LA	Least axis
LDE	Large dependence emphasis
LGLE	Low gray level emphasis
MTV	Metabolic tumor volume
NAC	Neoadjuvant chemotherapy
NGTDM	Neighboring gray tone difference matrix
PET-CT	Positron emission tomography-computed tomography
RE	Run entropy
RLNU	Run length nonuniformity
ROC	Receiver operating characteristic
SUVmax	Maximum standardized uptake value
SUVmean	Mean standardized uptake value
SVR	Surface volume ratio
SZNU	Size zone nonuniformity
TLG	Total lesion glycolysis
ZE	Zone entropy
ZP	Zone percentage
ZV	Zone variance

Introduction

Osteosarcoma is the most common primitive malignant neoplasm of the bone, which occurs most frequently in adolescents, with a second incidence peak among individuals aged > 60 years [1, 2]. The current treatment strategy—preoperative and postoperative adjuvant or neoadjuvant chemotherapy (NAC) in combination with definitive surgical resection—has dramatically improved the clinical survival rate from 20% to 60–70% compared with surgery alone [3]. Unfortunately, despite numerous clinical trials conducted over several decades, patient survival has not significantly improved. Osteosarcoma has markedly heterogeneous clinical behavior. Histopathological response to NAC, mainly the degree of necrosis, is reported to be the most important prognostic factor for survival in patients with localized osteosarcoma [4]. However, histological response cannot be evaluated for patient stratification before therapy. Currently, tumor necrosis can be assessed only in resected specimens after completion of NAC and continuation of ineffective chemotherapy can result in development of adverse effects, resistant clones, and associated costs [5]. Hence, to avoid ineffectual chemotherapy and optimize management of patients with localized osteosarcoma, an accurate and noninvasive means that can precisely predict the histological response to NAC is required in the era of personalized and precision medicine.

^{18}F-fluorodeoxyglucose (FDG) positron emission tomography-computed tomography (PET-CT) is a functional imaging modality and has been proposed as a noninvasive tool for predicting the chemotherapy response and clinical outcome in osteosarcoma [6–10]. Maximum standardized uptake value (SUVmax), mean standardized uptake value (SUVmean), metabolic tumor volume (MTV), and total lesion glycolysis (TLG), which are traditional parameters for analyzing ^{18}F-FDG PET data, are valuable factors in predicting tumor response and prognosis in patients with osteosarcoma [6–10]. Recently, it has been shown that tumor heterogeneity is an important factor correlated with aggressive behavior, decreased response to therapy, and disease progression [11–13]. Thus, radiomics, which includes texture analysis, has attracted increased attention. An approach has been proposed to quantify spatial tumor metabolic heterogeneity characterized by ^{18}F-FDG uptake texture features [14]. Increasing evidence shows that measurement of intratumoral heterogeneity on pretreatment ^{18}F-FDG PET images can give predictive information in several solid tumors, such as rectal [15], esophageal [16], and oropharyngeal cancer [17]. However, only a few studies have explored whether baseline ^{18}F-FDG PET radiomics features can predict treatment response or event-free survival (EFS) of patients with extremity osteosarcoma following NAC.

The aim of our study was to investigate whether the baseline texture features on ^{18}F-FDG PET images can predict treatment response to NAC or EFS in patients with localized extremity osteosarcoma.

Materials and methods

Patients, treatment, and follow-up

This retrospective cohort study was performed at a single institution from January 2013 to December 2017 at Shanghai Jiao Tong University Affiliated Sixth People's Hospital. The medical

records of patients with histologically confirmed and resectable high-grade (II or III) extremity osteosarcoma were studied. The inclusion criteria were as follows: (i) histologically proven primary high-grade extremity osteosarcoma; (ii) no history of treatment except biopsy; (iii) American Joint Committee on Cancer (AJCC) stage II; (iv) completion of NAC and adjuvant chemotherapy and surgery at our institute; (v) no more than 2 weeks between ^{18}F-FDG PET/CT and initiation of preoperative NAC; and (vi) follow-up for > 3 years for EFS. Exclusion criteria included the following: (i) patients with nonextremity osteosarcoma; (ii) secondary or extraskeletal osteosarcoma; (iii) skip lesions or distant metastases at initial presentation; (iv) incomplete clinical data; and (v) treatment by chemotherapy or surgery only. Thirty-five eligible patients were enrolled. Ethical approval was obtained for this single-institute retrospective analysis and written informed consent was waived.

All the eligible patients underwent two cycles of preoperative NAC followed by four cycles of postoperative chemotherapy. The preoperative NAC protocol was the modified T10 [18]. Definitive surgical resection of the primary tumor was scheduled between weeks 10 and 12 in the methotrexate-based protocol. After completion of treatment, routine follow-up evaluation was performed every 3 months for the first 2 years and every 6 months for the third year. In addition to clinical evaluation, each follow-up imaging examination included standard radiographs, CT scan, and magnetic resonance imaging of the operated limb. CT scans of the chest and whole-body bone scans were also performed.

^{18}F-FDG PET/CT imaging acquisition

Baseline ^{18}F-FDG PET/CT imaging was performed within 2 weeks of the patient commencing NAC. All patients were instructed to fast for at least 6 h before ^{18}F-FDG PET/CT. Plasma glucose level was determined before tracer injection, and a maximum value of 11 mmol/L was allowed. Image acquisition started 1 h after intravenous injection of ^{18}F-FDG at a dose of 3.7 MBq/kg. All studies were performed using GE Discovery VCT (General Electric Medical Systems) with the following setting: CT scan, 120 V and 80 mA, 64 slices, with a slice thickness of 3.75 mm. Whole-body PET scans were performed with a scan time of 2.5 min per bed position. Images were reconstructed iteratively using ordered subset expectation maximization. Attenuation correction was performed by unenhanced CT.

^{18}F-FDG PET/CT imaging analysis

For texture analysis, an experienced nuclear medicine physician (10 years' experience with ^{18}F-FDG PET/CT in musculoskeletal tumors) manually drew the region of interest (ROI) around the primary tumors on each transaxial slice solely on ^{18}F-FDG PET, not including CT data, while blinded to the patient outcome data. The ROIs were checked and validated by an independent senior nuclear medicine physician (22 years' experience). The segmentation was performed on ITK-SNAP (latest version 3.6.0) [19]. The ROIs of the

Fig. 1 Heat map depicting absolute value of Pearson correlation coefficients between all pairs of standard (SUVmax, SUVmean, MTV, and TLG) and texture PET features
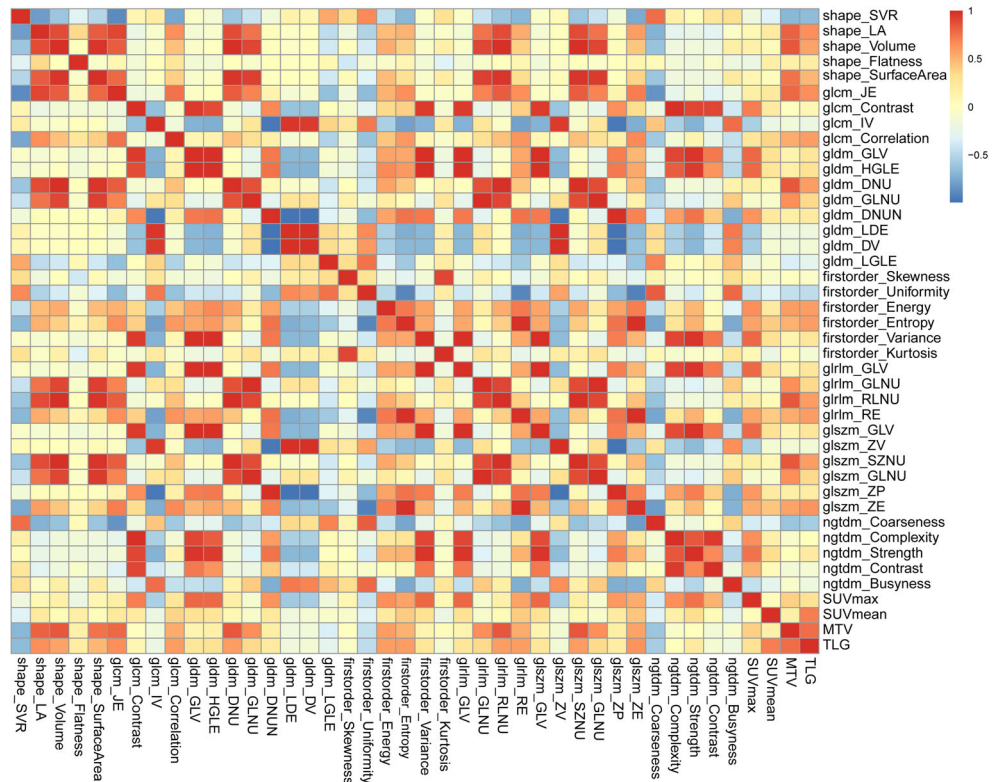
primary tumors were saved as DICOM-RT structures that were imported to the package of PyRadiomics imaging texture analysis toolbox (http://pyradiomics.readthedocs.io/en/latest/index.html) [20]. Specifically, 38 first- and higher-order texture features were derived and used for further analysis because these features are commonly used in medical imaging research [21]. SUVmax and SUVmean were defined as the maximum and mean tissue concentration in the delineated ROI, respectively. MTV was obtained by a threshold SUV of ≥ 2.0, which was used in osteosarcoma [6, 22]. TLG was calculated as (MTV) × (SUVmean).

Study outcomes

To improve the quality of our retrospective study, all pathological specimens were re-evaluated. The results of the reassessment were consistent with the initial pathological findings (completed by an expert musculoskeletal pathologist). Histological responses to NAC were graded as I and II, indicating a poor response, or grade III and IV, indicating a good response [23]. EFS was defined as the time interval from the first day of chemotherapy to local recurrence or distant metastasis, or to the appearance of secondary tumors or death from any cause, or to the last contact without events, whichever occurred first. Patients who were alive without recurrence or metastasis at the time of the last follow-up were censored.

Statistical analysis

Statistical analyses were performed using MedCalc software package (Version 18.2.1) and R open source statistical software (version 3.3.3; R project, http://www.r-project.org). As a preliminary step, to avoid feature redundancy, Pearson's correlations between all pairs of features were calculated and displayed via a heat map (Fig. 1). Features were reduced to form a signature of potentially uncorrelated features, but retaining those features that were associated with cancer biological characteristics and prognosis. According to the histological response, all ^{18}F-FDG PET features were compared using Student's t test for paired data. Those predicting better histological response ($p < 0.05$) were chosen, and areas under the receiver operating characteristic (ROC) curves (AUCs) were calculated for each parameter to determine the best predictor cutoff value in predicting histological response. Based on these cutoff values, patients were grouped and the performance of ^{18}F-FDG PET/CT was assessed as sensitivity, specificity, accuracy, positive predictive value, and negative predictive value. The Cox regression analysis was used to assess the effects of the texture parameters and other variables on EFS. To allow testing of multiple variables, all parameters with a p value < 0.01 were considered to indicate a significant difference in the univariate analysis and combined into a multivariate analysis to identify

those independently associated with EFS. Kaplan-Meier curves were calculated for parameters that showed statistical significance after multivariate analysis, and differences were assessed by using a log-rank test.

Results

Patient and tumor characteristics

The demographic and clinical characteristics of patients are detailed in Table 1. Between January 2013 and December 2017, 35 eligible patients were enrolled, of whom four had American Joint Committee on Cancer (AJCC) stage IIA osteosarcoma and 31 had stage IIB. The median age was 33 years and more than half the patients were male. The most frequent location of the primary tumor was the femur (49%), followed

Table 1 Clinical characteristics of the patients

Characteristic	Value
Age (years)	
Median	33
Range	7–76
Sex, n (%)	
Male	20 (57%)
Female	15 (43%)
Pathologic subtype, n (%)	
Osteoblastic	22 (63%)
Chondroblastic	7 (20%)
Others	6 (17%)
AJCC stage, n (%)	
IIA	4 (11%)
IIB	31 (89%)
Location of primary tumor, n (%)	
Femur	17 (49%)
Tibia	5 (14%)
Humerus	9 (26%)
Other	4 (11%)
Pathological fracture	
Yes	8 (23%)
No	27 (77%)
Alkaline phosphatase, n (%)	
Normal	20 (57%)
High	15 (43%)
Lactate dehydrogenase, n (%)	
Normal	26 (74%)
High	9 (26%)
Histological response, n (%)	
Good	16 (46%)
Poor	19 (54%)

by the tibia (14%). The histological subtype was osteoblastic in 22 patients (63%), chondroblastic in seven (20%), and others in six patients (17%). The alkaline phosphatase and lactate dehydrogenase levels were high in 15 and nine patients, respectively. Pathological findings revealed complete response in 16 patients and incomplete response in 19. The follow-up interval ranged from 38 to 52 months, with a median length of 41 months. No patients were lost. At the time of analysis, 20 (57%) patients had EFS, whereas the remaining 15 patients had experienced local recurrence, distant metastases, or death. Nine patients had metastatic disease: six in the lungs only, two in the bone only, and one in both the lungs and bone.

Tumor response prediction

Comparison of baseline ^{18}F-FDG PET parameters between responders ($n = 16$) and nonresponders ($n = 19$) is summarized in Supplemental Table 3. Neither SUVmax nor SUVmean was significantly different between responders and nonresponders. However, nonresponders showed significantly higher MTV and TLG. In texture analysis, 10 indices were different between the two groups. Nonresponders showed significantly higher least axis (LA), correlation, dependence nonuniformity (DNU), energy, entropy, run length nonuniformity (RLNU), run entropy (RE), and size zone nonuniformity (SZNU), but lower surface volume ratios (SVR) and coarseness. The AUCs for different parameters in predicting significant tumor nonresponsiveness were 0.918 for MTV, 0.862 for LA, 0.882 for DNU and SZNU, and 0.868 for RLNU. MTV was the most effective parameter for

predicting treatment response. Diagnostic performance for prediction of a poor histological response is summarized in Table 2 and Fig. 2.

EFS prediction

Univariate and multivariate Cox regression analyses were used to assess the effect of clinical parameters and ^{18}F-FDG PET variables on EFS (Table 3). In univariate analyses, many ^{18}F-FDG PET features were significantly correlated with EFS, but when combined in multivariate analysis, it modelled response (hazard ratio [HR] = 11.5029, $p = 0.0003$) + MTV (HR = 19.973, $p < 0.0001$) + coarseness$_{NGTDM}$ (neighboring gray tone difference matrix) (HR = 5.847, $p = 0.005$), which were significantly and independently associated with EFS. Kaplan-Meier curves were calculated (Fig. 3) and differences were assessed by log-rank test (Supplemental Table 4).

Discussion

In this context, we examined different baseline ^{18}F-FDG PET image-based parameters related to conventional and heterogeneity texture indices in patients with high-grade localized osteosarcoma of the extremities treated by NAC. The main findings indicated that measurements of texture features of ^{18}F-FDG uptake can predict treatment response (LA, DNU, SZNU, and RLNU) and EFS (coarseness$_{NGTDM}$), but the conventional parameter MTV provides the best predictive power.

Table 2 Areas under ROC curves for ability of ^{18}F-FDG PET pretreatment parameters to predict non-response

Classification of matrix	Features	AUC (95%CI)	Se (%)	Sp (%)	PPV (%)	NPV (%)	AC (%)
Conventional PET-related parameter	MTV	0.918 (0.774–0.983)	89.47	87.50	89.50	87.50	88.57
	TLG	0.839 (0.676–0.941)	89.47	81.25	85.00	86.70	85.71
Shape	SVR	0.842 (0.679–0.943)	73.68	87.50	87.50	73.70	74.29
	LA	0.862 (0.703–0.955)	68.42	93.75	92.90	71.40	80.00
GLCM	Correlation	0.717 (0.540–0.856)	94.74	43.75	66.70	87.50	68.57
GLDM	DNU	0.882 (0.727–0.966)	73.68	93.75	93.30	75.00	82.86
First order	Energy	0.717 (0.540–0.856)	78.95	56.25	68.20	69.20	57.14
	Entropy	0.684 (0.506–0.830)	63.16	68.75	70.60	61.10	62.86
GLRLM	RLNU	0.868 (0.711–0.958)	78.95	87.50	88.20	77.80	80.00
	RE	0.688 (0.509–0.833)	63.16	68.75	70.60	61.10	62.86
GLSZM	SZNU	0.882 (0.727–0.966)	73.68	93.75	93.30	75.00	82.86
NGTDM	Coarseness	0.836 (0.672–0.939)	63.16	100.00	100.00	69.60	80.00

Abbreviations: *GLCM*, gray level co-occurrence matrix; *GLDM*, gray level dependence matrix; *GLRLM*, gray level run length matrix; *GLSZM*, gray level size zone matrix; *NGTDM*, neighboring gray tone difference matrix; *MTV*, metabolic tumor volume; *TLG*, total lesion glycolysis; *SVR*, surface-volume ratio; *LA*, least axis; *IV*, inverse variance; *GLV*, gray level variance; *DNU*, dependence nonuniformity; *RLNU*, run length nonuniformity; *RE*, run entropy; *SZNU*, size zone nonuniformity; *ZP*, zone percentage; *SD*, standard deviation; *ROC*, receiver operating characteristic; *AUC*, area under curve; *Se*, sensitivity; *Sp*, specificity; *PPV*, positive predictive value; *NPV*, negative predictive value; *AC*, accuracy; *CI*, confidence interval
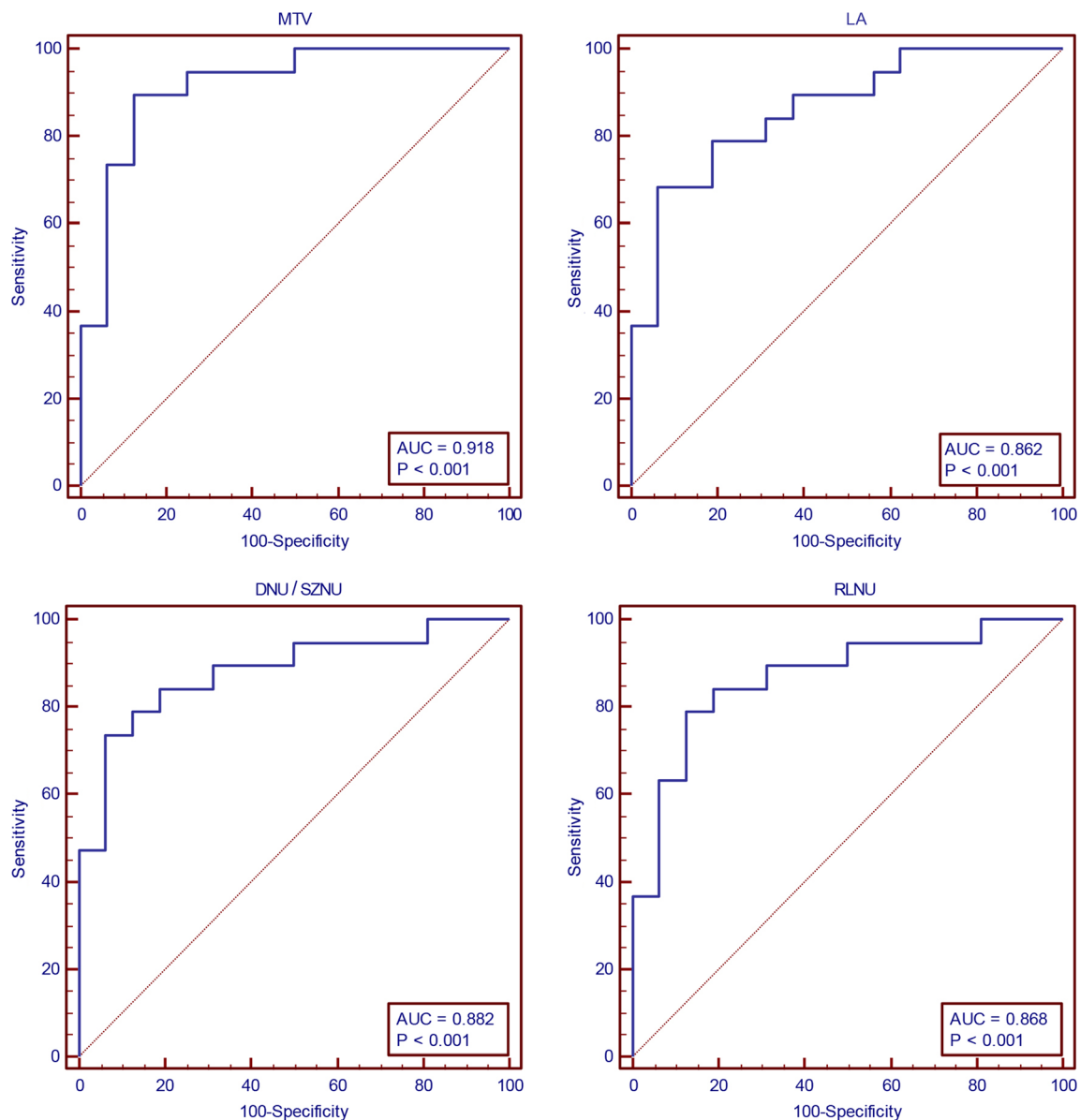
Fig. 2 ROC curves for baseline ^{18}F-FDG PET primary tumor MTV, LA, DNU, RLNU, and SZNU for identification of responders versus nonresponders. The area under the ROC curve for all above parameters was > 0.85

Tumor response prediction

Prediction of tumor response before treatment initiation is of value for optimizing treatment strategies for patients with osteosarcoma, especially for AJCC stage II. Patients with histologically low-grade osteosarcoma have a lower likelihood of metastases and are treated by surgery alone [24, 25]; patients with skip lesions and distant metastases at initial presentation have been reported to have poor outcomes [24]. Therefore, such patients were excluded from the current study.

Many studies have investigated the value of ^{18}F-FDG uptake in predicting treatment response in patients with osteosarcoma. As for baseline primary tumor uptake parameters to predict responders and nonresponders, MTV and TLG, but not

SUVmax and SUVmean, were significant discriminators in our study, especially MTV. Consistent with our data, it has been reported that initial MTV measured by ^{18}F-FDG PET can be used to predict the outcome of osteosarcoma of the extremities [6]. Byun et al [8] and Hyung et al [22] have revealed that the histological response to NAC in osteosarcoma can be predicted by MTV after one course of chemotherapy, rather than before treatment. MTV, as a conventional PET-related parameter, is based on metabolism of the malignant tumor and hence more accurately reflects the real tumor burden. It can be readily measured by drawing a VOI around the tumor, enabling high intra- and interobserver reproducibility [26]. It is also our clinical experience that MTV, rather than other PET imaging parameters, such as SUV, is more effective

Table 3 Cox proportional hazards model for EFS in patients with localized osteocarcoma

Classification of matrix	Features	HR (95%CI)	p value
Univariable analysis			
Clinical parameters	Age (years)	1.2665 (0.3570–4.4925)	0.7146
	Sex	0.6453 (0.2204–1.8896)	0.4243
	Pathologic type	0.8037 (0.2759–2.3621)	0.6959
	Location	1.2954 (0.7457–3.3346)	0.1121
	Fracture	1.3792 (0.4388–4.3350)	0.5821
	ALP	0.6771 (0.2312–1.9825)	0.4768
	LDH	1.2197 (0.3879–3.8345)	0.7340
	AJCC stage	3.2693 (0.8373–4.2097)	0.9531
	Response	8.4287 (1.8917–37.5538)	0.0052
Conventional PET parameters	SUVmax	2.3805 (0.8546–6.6312)	0.097
	SUVmean	3.5993 (1.1415–11.3487)	0.0288
	MTV	11.6998 (2.6164–52.3184)	0.0013
	TLG	14.0715 (1.8422–107.4847)	0.0108
Shape	SVR	5.5544 (1.8713–16.4865)	0.002
	LA	7.7816 (2.5853–23.4221)	0.0003
	Flatness	2.0101 (0.7280–5.5501)	0.1779
GLCM	Contrast	2.2152 (0.7557–6.4940)	0.1472
	IV	1.9788 (0.7138–5.4854)	0.1896
	Correlation	2.005 (0.7249–5.5459)	0.1802
GLDM	GLV	3.1098 (1.0978–8.8090)	0.0327
	DNU	8.5191 (2.6485–27.4024)	0.0003
First order	Skewness	2.0994 (0.4730–9.3176)	0.3293
	Uniformity	2.8774 (1.0192–8.1235)	0.0459
	Energy	5.4866 (1.9251–15.6372)	0.0014
	Entropy	2.9560 (1.0055–8.6905)	0.0488
	Kurtosis	1.6776 (0.6071–4.6352)	0.3184
GLRLM	RLNU	7.2902 (2.4265–21.9031)	0.0004
	RE	2.9560 (1.0055–8.6905)	0.0488
GLSZM	SZNU	6.8756 (2.1530–21.9575)	0.0011
	ZP	2.8144 (0.9551–8.2928)	0.0406
NGTDM	Coarseness	5.2418 (1.6490–16.6621)	0.0015
	Complexity	2.8315 (0.8983–8.9251)	0.0756
	Strength	1.8931 (0.6852–5.2301)	0.2183
	Contrast	4.0286 (1.3541–11.9853)	0.0122
	Busyness	2.5542 (0.5646–11.5562)	0.2234
Multivariable analysis			
Clinical parameter	Response	11.5029 (3.0429–42.6037)	0.0003
Conventional PET parameter	MTV	19.973 (4.048–98.533)	< 0.0001
NGTDM	Coarseness	5.847 (1.722–19.849)	0.005

Abbreviations: *HR*, hazard ratio; *CI*, confidence interval; others are the same as those for Table 2

in evaluating the therapeutic response and outcome of patients with osteosarcoma. In contrast to our findings, it has been demonstrated that SUVmax serves as a useful prognostic biomarker for osteosarcoma [10, 27, 28, and]. This may be because of the different inclusion criteria for our study and others. The other studies focused on pediatric and young adult patients or patients with different AJCC stages. However, Bailly et al [29] have demonstrated that SUVmax, MTV, and TLG were all not able to discriminate among histopathological responders. There are three possible explanations. First, the study focused on pediatric patients but with mixed AJCC stages, which was different from our study. Second, this team pooled osteosarcoma and Ewing's sarcoma. Third, the absence of sufficiently
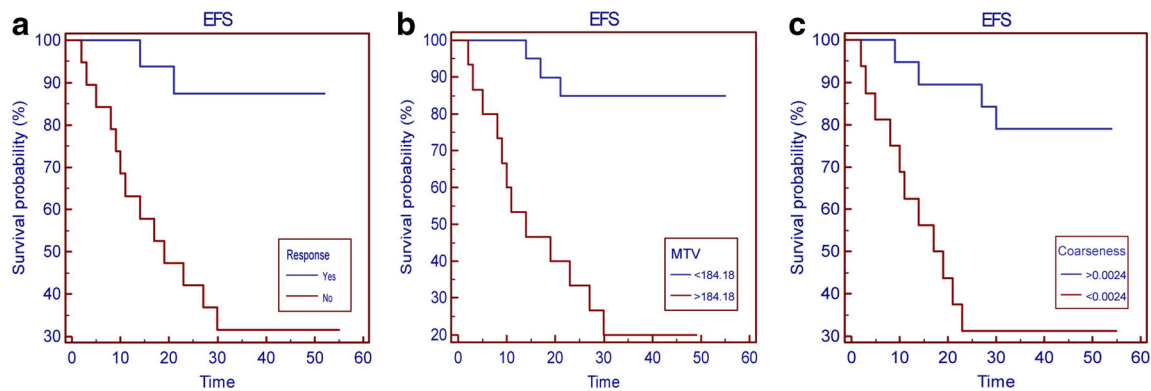
Fig. 3 Kaplan-Meier survival curves of the EFS demonstrate differences in patients in response (**a**), MTV (**b**), and coarseness$_{NGTDM}$ (**c**). The significant differences are 0.0008, 0.000,1 and 0.0018, respectively (log-rank test)

robust delineation techniques for tumor volume segmentation made consensus difficult.

All the above studies investigated only the conventional parameters of ^{18}F-FDG PET without texture indices, which are reported to play a crucial role in various solid malignancies. To the best of our knowledge, there are no data on ^{18}F-FDG PET texture analysis to predict tumor response of patients with extremity osteosarcoma treated by NAC. In our study, ^{18}F-FDG PET texture analysis results revealed that there were four texture features that had a better performance and could predict treatment response: LA, DNU, RLNU, and SZNU. For both LA and RLNU, the predictive accuracy for nonresponders was 80%, compared with 82.86% for DNU and SZNU. Our study is the first to indicate that the four features (LA, DNU, RLNU, and SZNU) are associated with treatment response of osteosarcoma. When compared with the traditional indicators, MTV provides the best predictive power, with an accuracy of 88.57%. The possible reasons were as follows. First, due to the rarity of osteosarcoma, the number of patients included in this study was small. Second, we only studied patients with AJCC stages IIA and IIB, but did not include other stages comprehensively. Third, we only studied PET images without comparing CT and MR images at the same time. Our study is only the first step, using texture analysis, in predicting the treatment response of osteosarcoma. Prospective studies incorporating more patients and in which the staging criteria are expand are still needed to confirm our results and to evaluate the predictive value of other texture analysis parameters, especially the newly discovered indicators DNU and SZNU.

Prognosis prediction

In our study, MTV, histological response, and the texture feature of coarseness$_{NGTDM}$ were independent predictors for EFS in the multivariate analysis. Consistent with our research, it has been demonstrated that MTV and histological response are independent predictors of metastasis in patients with osteosarcoma of the extremities [6]. Specifically, coarseness$_{NGTDM}$ was a newly

discovered texture feature, which was independently associated with EFS of osteosarcoma. The impact factor was derived from NGTDM, which included five fundamental parameters (coarseness, contrast, strength, busyness, and complexity). These parameters describe local tumor texture based on differences between each voxel and the neighboring voxels in adjacent image planes. Coarseness is a measure of average difference between the center voxel and its neighbors. It is linked to granularity within an image and is the most fundamental property of texture. A higher value indicates a lower spatial change rate and a locally more uniform texture. In esophageal carcinoma, coarseness is reported to discriminate responders to chemoradiotherapy from nonresponders [30]. Coarseness is also a feature that best differentiates head and neck cancer primary and nodal from normal tissues [31]. In non-small cell lung cancer, tumor coarseness is an independent predictor of overall survival in multivariate analysis [32]. Although texture features in functional imaging of cancer have potential clinical value, the relationship between biological characteristics of tumors and texture features is complex and largely unknown. Therefore, there is an urgent need to investigate thoroughly the texture features from different imaging modalities and to use different PET tracers to correlate with histopathological features that may influence patient stratification, treatment response, prognosis, and even gene expression.

Overall, our results add to the accumulating evidence that measurements of texture features of ^{18}F-FDG uptake within baseline PET images of osteosarcoma can predict treatment response (LA, DNU, SZNU, and RLNU) and EFS (coarseness$_{NGTDM}$). We consider these findings to be encouraging from the perspective of personalized medicine. By identifying patients with localized osteosarcoma of the extremities, who might respond poorly to NAC, oncologists can adjust the course of NAC to improve response. In contrast, patients with better prognosis could eventually benefit from less-aggressive treatment with the objective of reducing the toxicity of chemotherapy, as well as related costs.

Although promising, the results of this research should be interpreted cautiously. First, the retrospective nature of our

investigation means that we cannot exclude the possibility of selection bias. Second, due to the rarity of osteosarcoma and the homogeneous subgroups that we chose, the number of patients was small and in a single institution. Our ongoing prospective study will validate the predictive values of these parameters in a large number of patients. Third, we included only patients with localized high-grade extremity osteosarcoma, and thus, the conclusion of this study cannot be applied to patients with other AJCC stages. Finally, the radiomics study is constantly updating and developing, and further studies should include different methods for extracting texture features and determining the optimal approach.

Conclusions

Our present investigation suggests that intratumoral heterogeneity of baseline ^{18}F-FDG uptake measured by PET texture analysis can predict tumor response of patients with extremity osteosarcoma treated by NAC, but the conventional parameter MTV provides the best predictive power. MTV, histological response, and the texture feature of coarseness$_{NGTDM}$ were independently associated with EFS. The utility of such features, especially coarseness$_{NGTDM}$, should be confirmed by larger clinical studies before considering their potential integration into decisional algorithms aimed at personalized medicine.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Professor Quanyong Luo, email: lqyn@sh163.net.

Conflict of interest The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry Hongjun Song and Qian Wang have significant statistical expertise and no complex statistical methods were necessary for this paper.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Methodology
• retrospective
• diagnostic or prognostic study
• performed at one institution

References

1. ESMO/European Sarcoma Network Working Group (2014) Bone sarcomas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 25(Suppl 3):iii113–iii123
2. Mirabello L, Troisi RJ, Savage SA (2009) Osteosarcoma incidence and survival rates from 1973 to 2004: data from the Surveillance, Epidemiology, and End Results Program. Cancer 115:1531–1543
3. Bacci G, Longhi A, Fagioli F, Briccoli A, Versari M, Picci P (2005) Adjuvant and neoadjuvant chemotherapy for osteosarcoma of the extremities: 27 year experience at Rizzoli Institute, Italy. Eur J Cancer 42:2836–2845
4. Davis AM, Bell RS, Goodwin PJ (1994) Prognostic factors in osteosarcoma: a critical review. J Clin Oncol 12:423–431
5. Bajpai J, Gamnagatti S, Kumar R et al (2011) Role of MRI in osteosarcoma for evaluation and prediction of chemotherapy response: correlation with histological necrosis. Pediatr Radiol 41:441–450
6. Byun BH, Kong CB, Park J et al (2013) Initial metabolic tumor volume measured by ^{18}F-FDG PET/CT can predict the outcome of osteosarcoma of the extremities. J Nucl Med 54:1725–1732
7. Byun BH, Kong CB, Lim I et al (2013) Combination of 18F-FDG PET/CT and diffusion-weighted MR imaging as a predictor of histologic response to neoadjuvant chemotherapy: preliminary results in osteosarcoma. J Nucl Med 54:1053–1059
8. Byun BH, Kong CB, Lim I et al (2014) Early response monitoring to neoadjuvant chemotherapy in osteosarcoma using sequential ^{18}F-FDG PET/CT and MRI. Eur J Nucl Med Mol Imaging 41:1553–1562
9. Palmerini E, Colangeli M, Nanni C et al (2017) The role of FDG PET/CT in patients treated with neoadjuvant chemotherapy for localized bone sarcomas. Eur J Nucl Med Mol Imaging 44:215–223
10. Davis JC, Daw NC, Navid F et al (2018) 18F-FDG uptake during early adjuvant chemotherapy predicts histologic response in pediatric and young adult patients with osteosarcoma. J Nucl Med 59:25–30
11. Rajendran JG, Schwartz DL, O'Sullivan J et al (2006) Tumor hypoxia imaging with [F-18] fluoromisonidazole positron emission tomography in head and neck cancer. Clin Cancer Res 12:5435–5441
12. van Velden FH, Cheebsumon P, Yaqub M et al (2011) Evaluation of a cumulative SUV-volume histogram method for parameterizing heterogeneous intratumoural FDG uptake in non-small cell lung cancer PET studies. Eur J Nucl Med Mol Imaging 38:1636–1647
13. Asselin MC, O'Connor JP, Boellaard R et al (2012) Quantifying heterogeneity in human tumours using MRI and PET. 48:447–455
14. Cook GJ, O'Brien ME, Siddique M et al (2015) Non-small cell lung cancer treated with erlotinib: heterogeneity of (18)F-FDG uptake at PET-association with treatment response and prognosis. Radiology 276:883–893
15. Lovinfosse P, Polus M, Van Daele D et al (2018) FDG PET/CT radiomics for predicting the outcome of locally advanced rectalcancer. Eur J Nucl Med Mol Imaging 45:365–375
16. Beukinga RJ, Hulshoff JB, van Dijk LV et al (2017) Predicting response to neoadjuvant chemoradiotherapy in esophageal cancer with textural features derived from pretreatment (18) F-FDG PET/CT imaging. J Nucl Med 58:723–729
17. Cheng NM, Fang YH, Lee LY et al (2015) Zone-size nonuniformity of 18F-FDG PET regional textural features predicts survival in patients with oropharyngeal cancer. Eur J Nucl Med Mol Imaging 42:419–428
18. Rosen G, Marcove RC, Huvos AG et al (1983) Primary osteogenic sarcoma: eight-year experience with adjuvant chemotherapy. J Cancer Res Clin Oncol 106(Suppl):55–67

19. Yushkevich PA, Piven J, Hazlett HC et al (2006) User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 31:1116–1128

20. van Griethuysen JJM, Fedorov A, Parmar C et al (2017) Computational radiomics system to decode the radiographic phenotype. Cancer Res 77:e104–e107

21. Chicklore S, Goh V, Siddique M, Roy A, Marsden PK, Cook GJ (2013) Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis. Eur J Nucl Med Mol Imaging 40:133–140

22. Im HJ, Kim TS, Park SY et al (2012) Prediction of tumour necrosis fractions using metabolic and volumetric 18F-FDG PET/CT indices, after one course and at the completion of neoadjuvant chemotherapy, in children and young adults with osteosarcoma. Eur J Nucl Med Mol Imaging 39:39–49

23. Coffin CM, Lowichik A, Zhou H (2005) Treatment effects in pediatric soft tissue and bone tumors: practical considerations for the pathologist. Am J Clin Pathol 123:75–90 Review

24. Greene FL (2002) AJCC cancer staging manual, 6th edn. Springer-Verlag, New York, pp 187–192

25. Carrle D, Bielack SS (2006) Current strategies of chemotherapy in osteosarcoma. Int Orthop 30:445–451

26. Chen HH, Chiu NT, Su WC, Guo HR, Lee BF (2012) Prognostic value of whole-body total lesion glycolysis at pretreatment FDG PET/CT in non-small cell lung cancer. Radiology 264:559–566

27. Kong CB, Byun BH, Lim I et al (2013) 18F-FDG PET SUVmax as an indicator of histopathologic response after neoadjuvant chemotherapy in extremity osteosarcoma. Eur J Nucl Med Mol Imaging 40:728–736

28. Im HJ, Zhang Y, Wu H et al (2018) Prognostic value of metabolic and volumetric parameters of FDG PET in pediatric osteosarcoma: a hypothesis-generating study. Radiology 287:303–312

29. Bailly C, Leforestier R, Campion L et al (2017) Prognostic value of FDG-PET indices for the assessment of histological response to neoadjuvant chemotherapy and outcome in pediatric patients with Ewing sarcoma and osteosarcoma. PLoS One 12:e0183841

30. Tixier F, Le Rest CC, Hatt M et al (2011) Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. J Nucl Med 52:369–378

31. Yu H, Caldwell C, Mah K et al (2009) Coregistered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning. IEEE Trans Med Imaging 28:374–383

32. Cook GJ, Yip C, Siddique M et al (2013) Are pretreatment 18F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? J Nucl Med 54:19–26